DOCUMENT RESUME

ED 479 692 SO 035 149

AUTHOR	Thorpe, Pamela K.
TITLE	Escher's Intersecting Worlds: Evaluation as a Reflection of the Evaluator, The Evaluator Being Reflected in the Evaluation.
PUB DATE	2003-00-00
NOTE	12p.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Evaluation Methods; *Literacy; Middle Schools; Program Evaluation; Public Schools; *Reading Comprehension; *Reading Programs; Self Evaluation (Individuals)
IDENTIFIERS	*Escher (M C)

ABSTRACT

A program evaluation (of Scholastic's Read 180 program) was recently submitted to administrators at the public middle school where the author/educator teaches. The evaluation was written in a reflective state regarding the nature of evaluations and her role in the construction of this particular evaluation. It recalls the M. C. Escher lithograph, "Drawing Hands," which portrays hands drawing themselves, thereby conveying the message that self and self-reference are indivisible and coequal. The evaluation of an intervention program, with names of schools removed, is presented in this paper. The thoughts regarding the potential bias affecting the evaluation, and the competencies required to produce the evaluation are presented. She notes that she felt like a character in a book, albeit a shadowy character, reading her own story in the midst of this evaluation. (Author/BT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 479 692

Escher's Intersecting Worlds: Evaluation as a Reflection of the Evaluator, the Evaluator being Reflected in the Evaluation

Pamela K. Thorpe, PhD.
Wichita Public Schools
Wichita, Kansas

SO 035 149

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

P. K. Thorpe

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

BEST COPY AVAILABLE

2

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Abstract

A program evaluation was recently submitted to administrators at our public school. I, as the writer of this evaluation, waxed into a reflective state regarding the nature of evaluations and my role in the construction of this particular evaluation. The evaluation of a intervention program, with names of schools removed, is presented first in this paper. The author's thoughts regarding the potential bias effecting the evaluation, and the competencies required to produce the evaluation is presented next. I indeed felt like a character in a book, albeit a shadowy character, reading my own story in the midst of this evaluation.

Escher's Intersecting Worlds: Evaluation as a Reflection of the Evaluator, the Evaluator being Reflected in the Evaluation

The lithograph, "Drawing Hands" by M. C. Escher (1948), portrays hands drawing themselves, thereby conveying the message that self and self-reference are indivisible and co-equal. Self-reference may be found in the way our worlds of perception reflect and intersect one another. We seem like a character in a book that is reading his/her own story. "Drawing Hands", therefore, is a work reflecting the artist and the artist being reflected in his work. Similarly, I believe that a program evaluation reflects the evaluator with the evaluator being reflected in the evaluation. As esoteric as this sounds, it is important to realize that a program evaluation may reflect the biases and competencies of the evaluator. Consequently, upon close examination, the evaluator is indeed being reflected in her/his evaluation.

An evaluation of Scholastic, Inc.'s Read 180 program was recently submitted to administrators at our public school. I, as the writer of this evaluation, waxed into a reflective state regarding the nature of evaluations and my role in the construction of this particular evaluation. The evaluation of the Read 180 program, with names of schools removed, is presented first in this paper. The author's thoughts regarding the potential bias effecting the evaluation and the competencies required to produce the evaluation will then be presented. I indeed felt like a character in a book, albeit a shadowy character, reading my own story in the midst of this evaluation.

The Evaluation

Read 180 First Year Evaluation (2002-2003)

Data-driven, or rather data-informed decision-making is foremost on every educator's mind as we embark on the era of No Child Left Behind. The first-year evaluation of Scholastic's Read 180 program is designed to aid in that data-informed decision-making process. Therefore, this evaluation begins with a brief review of one research article regarding computer-assisted reading instruction across multiple programs, and examines one evaluation study of the Read 180 intervention program in particular.

The Mid-continent Research for Education and Learning (MCREL) regional educational laboratory recently published a report entitled, "Helping At-Risk Students Meet Standards: A Synthesis of Evidence-Based Classroom Practices" (Barley, Lauer, Arens, Apthorp, Englert, Snow & Akiba, 2002). This document addresses the question, "What are effective instructional strategies that can be used in classrooms to assist low-achieving students?" Researchers and teachers have for decades believed in the efficacy of computer-based instruction for learning mathematics and enhancing literacy competencies. Accordingly, one chapter in MCREL's document examines the research literature on computer-assisted instruction. MCREL's meta-analysis of thirty-five research studies on literacy found no significant effect of computer-assisted literacy

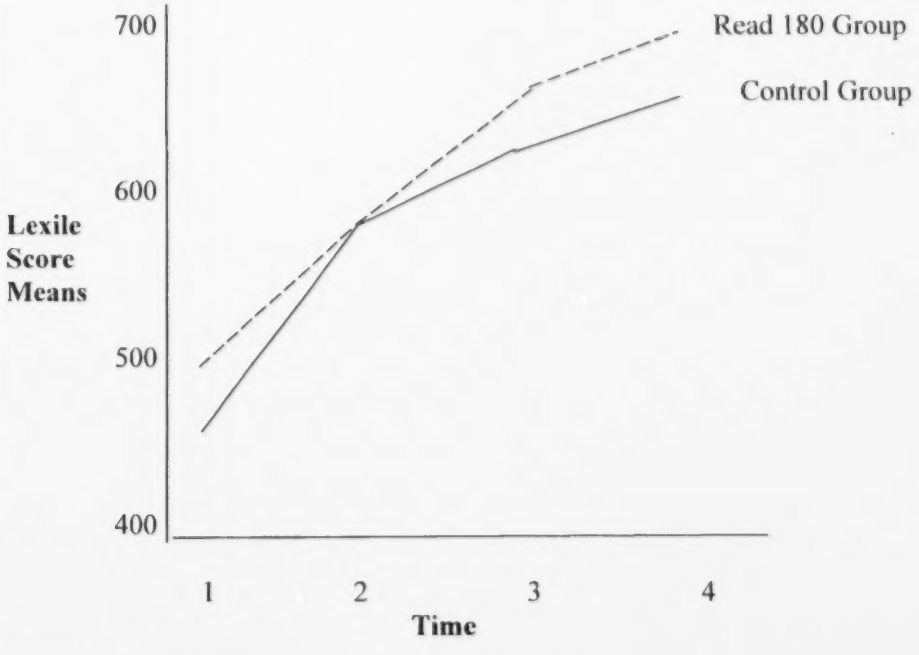
programs for enhancing the competencies of at-risk students. The effect size for these studies was .16 with a standard deviation of .40, leading to a confidence interval that included the value zero. A confidence interval that includes the value zero tells us that we find no support for the hypothesis that computer-assisted instruction for literacy is an effective strategy for increasing the literacy skills of low-achieving students over other instructional strategies. (Note: Effect size is a name given to a family of indices that measure the magnitude of a treatment effect. Unlike significance tests, these indices are independent of sample size. Effect size measures are the common currency of meta-analytic studies that quantitatively summarize findings from a specific area of research.) It is important to keep their findings in mind as we examine the potential of Read 180 for improving the reading competencies of low-achieving children.

An examination of Scholastic's research publications regarding the Read 180 program is also warranted. We need to understand the likely effect this particular reading intervention has towards helping children to obtain reading competencies. Becker, Mann & Sweeney (2001) presented findings from a validation study regarding Read 180 and The Council of Great City Schools. This final report was written by Interactive, Inc. who was contracted to conduct an independent validation study of the effects of Read 180 for low performing students in: Atlanta, GA; Boston, MA; Columbus, OH; Dallas, TX; Houston, TX; Miami-Dade, FL; and San Francisco, CA. Read 180 students, from three school districts (Boston, Dallas and Houston), which had year-to-year scores on the Stanford-9 (Total Reading), showed a significant difference in growth (Mean = 22.94) over the control group (Mean = 17.24). Moreover, an analysis of covariance on the post-test Stanford-9 scores, controlling for Stanford-9 pre-test scores, showed a significant difference in favor of students who had been enrolled in the Read 180 program.

Upon further examination, it is apparent that the treatment and control groups were non-equivalent at the start of the study. The school districts had agreed to randomly assign students to the Read 180 treatment group and to a control group; they chose not to carry out this aspect of the evaluation. This is unfortunate since the statistical analyses presented in this validation study cannot compensate for their adjustment in research methodology. The gain score results do not provide adequate evidence that it is the Read 180 program that provides superior results over the control group. We simply cannot be sure that students in the Read 180 program did not have better reading knowledge and skills than the control group students at the start of the study, nor that they were equivalent on any other characteristics which might aid in the acquisition of reading comprehension skills. Furthermore, using analysis of covariance to statistically equate non-equivalent groups at the start of the study, though tempting and frequently used in educational studies, has been shown since the 1970's to be inadequate to the task, and thus should not be used for this purpose. Analysis of covariance may be appropriately used for enhancing power, i.e. the ability to find a treatment effect when there really is one, in studies with random assignment of students to treatment and control groups. Finally, even if we accept the findings of the analysis of covariance as appropriate, the effect sizes reported in this study are very small. For 6th grades in Boston, the effect size is.04, implying that only 4% of the variability in Stanford-9 Total Reading scores can be attributed to being in the Read 180 intervention program or in the control group. For the Houston 7th graders it is only 3%, for the Houston 8th graders it is 1% and for the Dallas 8th grade students it is less than 1%. Depressingly, these effects are also inflated in value

since they are the sample effects for this study and not the effects one can expect to find for the populations of 6th, 7th and 8th grade students.

We now turn to first year data collected from the Read 180 programs at School 1, School 2 and School 3 middle schools in the mid-western school district. The initial set of data being examined is for School 1 7th graders. A Read 180 treatment group and a control group were established by school administration. An inspection of the graphical display of test means from four time points across the school year is helpful at this point. These values were obtained from a multivariate repeated measures analysis with one factor being time (a within-subjects factor) and the other factor being group membership (a between-subjects factor denoting membership in the control group or the Read 180 group). Lexile scores, the measure of reading comprehension used by Read 180, is the dependent variable of interest for this evaluation.



Notice that Read 180 students begin with higher lexile scores on average than control group students. Findings from the analysis show a significant linear increase in lexile scores for both groups of students. There is also a significant quadratic component in the data, which can be seen in the trends as both groups of students begin to decline in their growth in lexile scores towards the end of the school year. The effect size for the linear aspect of growth is .43 and an effect size of .21 exists for a quadratic component of growth. More importantly, no significant group differences were found for linear and quadratic growth patterns. The Read 180 and control groups were growing at the same rate and slowing down in growth at the same rate. The group effect size is .004, meaning that less than 1% of the variability found in the growth of students is due to being in the

Read 180 classroom or in the control classroom. Thus, group membership does not explain students' lexile scores.

The conclusion from the analysis is that students on average have grown in reading comprehension, with a slowing growth rate during the last nine weeks of school. In addition, this growth pattern is equivalent for both the Read 180 students and the control group students.

The first year data for 7th graders at School 1 provides us with one picture of the Read 180 reading intervention program. It is also worthwhile to look at the relationship between Read 180 students' lexile scores and their reading comprehension scores as measured by the mid-western's public schools' Benchmark Assessment of 8th Grade Reading. This instrument is primarily an assessment of reading comprehension. Therefore, we will now examine the relationship of Read 180 8th grade students' lexile scores with their scores on the 8th Grade Benchmark Reading Assessment.

Data for 8th grade students at School 2 and School 3 was examined. No relationship was found between students' lexile scores (actual lexile scores across four time points, last lexile score taken, and growth scores) and their 8th Grade Benchmark Reading Assessment scores.

To summarize, the effectiveness of Read 180 over other methods for increasing the reading comprehension competencies of low-achieving children was not established by the first-year intervention data from School 1. Furthermore, no evidence was found relating students' lexile scores to the 8th Grade Reading Benchmark Assessment at School 2 and School 3.

Clear-cut conclusions regarding the effectiveness of the Read 180 program, based upon the research articles presented and the context of the first-year data cannot be supported from a data-informed decision-making perspective. Firstly, MCREL's meta-analysis included studies where the teacher's role depended upon the computer software used in the intervention, i.e. some software required minimal assistance from the teacher whereas others required teachers' input to facilitate instruction. Studies used in the meta-analysis may not accurately match the teacher's role in the Read 180 intervention program. In addition, the literacy test scores incorporated in the meta-analysis included not only reading, but also vocabulary, writing, listening and language scores. Therefore, though MCREL's analysis is useful in providing a context for our evaluation of the Read 180 program, it should in no way be viewed as deterministic regarding the outcome of Read 180 for enhancing the reading comprehension of low-achieving children. Secondly, data from the first-year implementation of the Read 180 program reflected the variability found across schools in their ability to actually begin to identify and work with students in Read 180. For instance, School 3 had to wait until after the start of the school year for their classroom to be ready for instruction. In addition, use of the 90-minute instructional time also varied across schools, i.e. School 2 utilized small group times for writing and other tasks. It seems that School 2 had school goals for teaching skills and concepts, and the School 2 teacher was attempting to meet those school goals during small group instruction. Finally, teachers had varying capabilities for implementing and utilizing all aspects of the Read 180 intervention program towards enhancing the reading comprehension capabilities of their students. Consequently, more data will need to be collected from students in the Read 180 program this coming year in order to make reliable decisions regarding the effectiveness of this particular reading intervention.

An Evaluation of the Evaluation and Evaluator

My recent move from a university faculty position to a position as a research analyst in a public school system has made me more conscious of No Child Left Behind's impact on public PreK to 12 educational institutions. This new era endorses the Liberal's objectives for public education in the framework of the Conservative's demand for tax-conscious, market-driven public institutions. Therefore, this evaluator, while writing the evaluation for Read 180, reflected upon the strong undercurrents of potentially adversarial philosophies in her, if not in the school district. Brandt (2003) eloquently provides us with a perspective for linking political, economic, and educational philosophies regarding literacy, an appropriate topic with respect to the Read 180 evaluation. Brandt says,

> "Literacy is changing because the economy is changing. The United States has become a so-called knowledge economy or informational economy, in which mental labor has replaced physical labor and making information and ideas has replaced making things as our main economic pursuit. Human capital is now regarded as more valuable than land or even money, so literacy has become a hot commodity." (p. 245)

Changes in our economy have indeed raised expectations for literacy achievement with a desire for a more equal distribution of literate skills within and between groups in our society. Brandt contends that this equal distribution is not happening; instead the increasing value of literacy is leading to greater ethnic and class inequity (Brandt, 2001). Her 2003 article addresses the question, "What does it mean to be a nation where literacy is taught and learned under the banner of economic productivity and competition?" Brandt contextualizes a response to this question by examining 'sponsors' of literacy in American lives. She contends that sponsors of literacy proliferate in the United States using the development of reading and writing skills to their own economic advantage, and consequently have an effect on people as students, parents, workers and citizens. I would also add that this has an effect on educators (teachers and administrators), and more particularly, it has had an effect on me. I find it difficult to believe most marketing information provided to educators by sponsors of literacy products. I tend to have a negative bias against their products, believing that results and conclusions for their products are at best incomplete and at worst deliberately distorted. Therefore, the objectivity I need for a thorough and honest evaluation is constantly under attack by this bias. The best I can hope to accomplish is to acknowledge this bias and guard against it by carefully scrutinizing the design of the research project and the conclusions drawn from the data. I have also found it helpful to allow an independent reader access to the design of the study and subsequently to the conclusions drawn from the study. Opening up the current evaluation of the Read 180 program to a broader educational audience at a national conference is also an attempt to keep the evaluation from being tainted by my own biases and limited by my competencies. It is to these competencies that I now turn.

My graduate training is as a developmental psychologist, with a strong emphasis in cognitive development, research methodology and statistical analyses used in longitudinal research activities (See Thorpe, 2003a). I have used the knowledge and skills from this training to address research questions regarding student learning in a classroom context. Figure 1, entitled an "Ecological Perspective of Children's Development in School", is my first attempt at conceptualizing my developmental psychology preparation in the context of a public school system (Thorpe, 2003b). This figure helps me to frame the research and evaluation projects for the district in terms of my professional preparation and with respect to the school system's organization and function within the larger context of our American culture. For instance, a child's cognitive development regarding symbol systems and the contexts within which that development takes place is found in the first dimension of the cube. This dimension signifies a child's physical, socio-emotional, and cognitive development. The second dimension, denoting the support systems for learning at school, includes relational, academic and developmental support. The final dimension in the cube acknowledges the structural importance of schools for student learning, i.e. resources, organization and leadership. The leadership in our school district recognizes the importance of increasing the literacy skills of low achieving middle school students. Therefore they have set aside monetary resources to pay for the Read 180 intervention in the three middle schools, and school principals have organized the school day and students' class schedules to accommodate Read 180's 90-minute requirements for their intervention. These decisions are on the one hand a response to the pressures applied from the change in political ideology at the federal level and it's subsequent No Child Left Behind legislation, but their decisions also reflect a sincere desire within administrators and teachers to enable children to become effective readers.

Preparation I received in research methodology and statistics also affects the way I review the literature regarding a particular research topic, or in this case the topic of this evaluation. Maxwell & Delaney (2003) clearly present their case against using ANCOVA to equate groups in non-randomized studies. They punctuate their position for not using ANCOVA to equate groups that differ in various ways by arguing that such an adjustment potentially diminishes differences in one dimension while increasing differences in another dimension. They add a quote from Lord (1967),

> With the data usually available for such studies, there is simply no logical or statistical procedure that can be counted on to make proper allowances for uncontrolled pre-existing differences between groups. p. 307

With respect to the current evaluation, it cannot be overlooked that the students at School 1 were not randomly assigned to receive the Read 180 intervention or to a control group. That the groups did not statistically differ from start to finish was fortunate. Otherwise a group difference in favor of the Read 180 group could not be clearly interpreted to be a result of the intervention itself. Support for a causal effect of Read 180 can only be derived from the nature of the research design and not from the statistical model. Statistical decisions are basically organized arguments, and are related to how experiments or evaluations are designed. Careful design of experiments and program evaluations, especially in our age of data-informed decision-making, have as their goal

sound inferences that can be fully justified and logically compelled by the data (Maxwell & Delaney, 2003). Continually working to create sound designs is essential if school administrators are to have reliable information for decision-making purposes.

In conclusion, I see my preparation as an asset to my current position in the school district, adding a set of knowledge and skills unique to others in the system. I am certainly aware that my background biases and competencies are reflected in the evaluation of the Read 180 program. It is yet to be seen if these biases can be contained, and my competencies be effectively applied to create evaluations that are useful for data-informed decision-making. The jury is still out on this question and I eagerly await the answer. Until then, and in the spirit of Escher, I encourage you to evaluate your evaluation of the evaluator in this evaluation.

References

Barley, Z., Lauer, P. A., Arens, S. A., Apthorp, H. S., Englert, K. S., Snow, D. & Akiba, M. (2002). Helping at-risk students meet standards: A synthesis of evidence-based classroom practices. Aurora, CO: Mid-continent Research for Education and Learning.

Becker, J., Shakeshaft, C., Mann, D., & Sweeney, K. (2002). Validation study: Final report for Scholastic Inc. and the Council of Great City Schools. Interactive, Inc.

Brandt, D. (2003). Changing literacy. Teachers College Record, 105(2), 245-260.

Brandt, D. (2001). Literacy in American lives. New York: Cambridge University Press.

Bronfenbrenner, R. (1998). The ecology of developmental processes. In R. M. Lerner (Ed.), Handbook of Child Psychology (5th Ed., Vol 1): Theory. [Series Editor: W. Damon], pp. 993-1028. NY: Wiley.

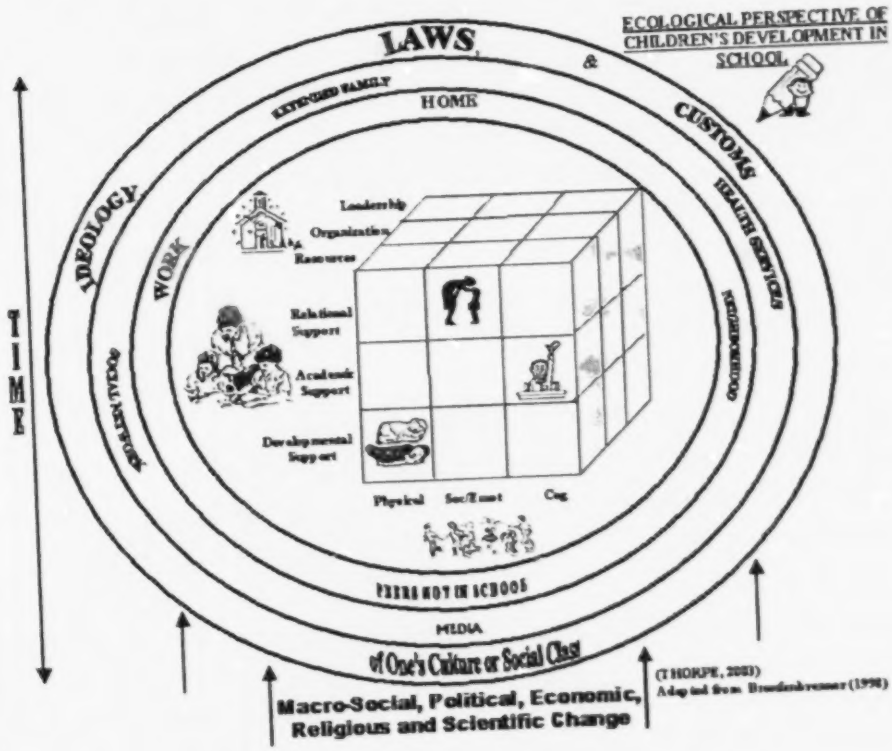
Lord, F. M. (1967). A paradox in the interpretation of group comparisons. Psychological Bulletin, 68, 304-305.

Maxwell, S. E. & Delaney, H. D. (2003). Designing experiments and analyzing data: A model comparison perspective (2nd Ed.). NJ: Lawrence Erlbaum Associates.

Thorpe, Pamela K. (2003a). Modeling children's academic development at school – Part I: Contrasting quantitative approaches to the analysis of change. Manuscript submitted for publication to ERIC.

Thorpe, Pamela K. (2003b). A mediation model relating teacher ratings of student achievement to students' connectedness at school. Paper presented at a paper discussion session for the American Educational Research Association, Chicago, IL.

Figure 1. Ecological perspective of children's development in school.



BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC Educational Resources Information Center

Reproduction Release

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Escher's Intersecting Worlds: Evaluation as a reflection of the evaluator, the evaluator being reflected in the evaluation.	
Author(s): Pamela K. Thorpe, Ph.D.	
Corporate Source: Wichita Public Schools, Wichita, KS	Publication Date: 7/2/03 (Submitted)

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
Level 1	**Level 2A**	**Level 2B**
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Pamela K. Thorpe	Printed Name/Position/Title: Pamela K. Thorpe, Research Analyst	
Organization/Address: Wichita Public Schools, 201 N. Water, Suite 316, Wichita, KS 67202	Telephone: 316-973-4727	Fax: 316-973-4730
	E-mail Address: pthorpe@usd259.net	Date: 7/2/03

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC/CHESS
2805 E. Tenth Street, #120
Bloomington, IN 47408